

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated:

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.

- If you encounter an accession number from an older search run against UniProt (results file extension .rnp) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any SIFC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2006, 14:24:44 ; Search time 234 Seconds

(without alignments)
1320.604 Million cell updates/sec

Title: US-10-788-746-6

Perfect score: 2316

Sequence: 1 AVSTSPAAFYDTQTLLAAHAA.....NPPPLTALTAACCPGARA 438

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 05.80:*
1: uniprot_sprot:*
2: uniprot_crembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	DB ID	Description
1	2316	100.0	1284 2	Q9WJ35_VIRU
2	2316	100.0	1320 2	Q8WJ35_VIRU
3	2234	96.5	1319 2	Q5JAM4_VIRU
4	2225	96.1	1319 2	Q5JAM6_VIRU
5	2224	96.0	1319 2	Q5JAM9_VIRU
6	2203	95.1	1322 2	Q8Q152_VIRU
7	2203	95.1	1324 2	Q9Q1V0_VIRU
8	859	37.1	167 2	Q9YQD9_VIRU
9	501	21.6	1258 2	Q8Q827_MIDDV
10	499	21.5	1244 2	Q8Q828_VIRU
11	487	21.0	1241 2	Q5WQY5_VIRU
12	477	20.6	1241 2	Q8Q821_VIRU
13	473	20.4	1241 2	Q8Q825_VIRU
14	471	20.3	1246 2	Q5KXP2_VIRU
15	470	20.3	1247 2	Q5KXP3_VIRU
16	470	20.3	1248 2	Q5KXP4_VIRU
17	468	20.2	1248 2	Q8JUX5_VIRU
18	465.5	20.1	1253 2	Q5Y388_VIRU
19	465	20.1	1247 2	Q9Q371_VIRU
20	462.5	20.0	1253 2	Q8Q831_VIRU
21	458.5	19.8	1231 2	Q8Q847_VIRU
22	456.5	19.7	1241 2	Q66579_EBEV
23	454.5	19.6	1240 1	POL5_EBEV
24	452.5	19.5	1239 2	P89946_BFV
25	452.5	19.5	1239 2	Q8Q823_BFV
26	452.5	19.5	1242 2	Q88795_EBEV
27	452.5	19.5	1242 2	Q88792_EBEV
28	452.5	19.5	1242 2	Q88790_EBEV
29	452.5	19.5	1242 2	Q88799_EBEV
30	452.5	19.5	1242 2	Q9PZX6_EBEV
31	452.5	19.5	1242 2	Q88798_EBEV

32	452.5	19.5	1242 2	Q88678_EBEV	Q88678 eastern equ
33	452.5	19.5	1242 2	Q4QXJ9_EBEV	Q4QXJ9 eastern equ
34	452.5	19.5	1242 2	Q4QXJ7_EBEV	Q4QXJ7 eastern equ
35	452	19.5	1247 1	POL5_ONNVG	P22056 O'nyong-nyo
36	450.5	19.5	1242 2	Q88754_EBEV	Q88754 eastern equ
37	450.5	19.5	1242 2	Q88793_EBEV	Q88793 eastern equ
38	449.5	19.4	1242 2	Q88796_EBEV	Q88796 eastern equ
39	449.5	19.4	1242 2	Q88797_EBEV	Q88797 eastern equ
40	447.5	19.3	1253 2	Q8Q833_VIRU	Q8Q833 sagiyama vi
41	446.5	19.3	1242 2	Q9PZX7_EBEV	Q9PZX7 eastern equ
42	445.5	19.2	1242 2	Q9PZX1_EBEV	Q9PZX1 eastern equ
43	445.5	19.2	1253 2	Q9PZX6_VIRU	Q9PZX6 sagiyama vi
44	442.5	19.1	1242 2	Q9PZX5_EBEV	Q9PZX5 eastern equ
45	442.5	19.1	1242 2	Q08359_EBEV	Q08359 eastern equ

ALIGNMENTS

Query Match	Best Local Similarity	100.0%	Score 2316	DB 2	Length 1284	Pred. No. 1.8e-171	Matches 438	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	AVSTSPAAFYDTQTLLAAHAAAPYRANCPCDGTACISPIAIDVSSGSHVLMRNGS	60								
DB	354	AVSTSPAAFYDTQTLLAAHAAAPYRANCPCDGTACISPIAIDVSSGSHVLMRNGS	413								
QY	61	OSGYTAGGAGGAGSTLYRGDGVHAADNTRLVRTAKCDVQATGHYTLIANGPCVGS	120								
DB	414	OSGYTAGGAGGAGSTLYRGDGVHAADNTRLVRTAKCDVQATGHYTLIANGPCVGS	473								
QY	121	LTVAATDGTGHOCTTVFHHQVTEKFTREBSKGNHLSMTKCTKTFSTTPKSAIYLVYD	180								

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2006, 14:31:58 ; Search time 48 Seconds
(without alignments)
754.416 Million cell updates/sec

Title: US-10-788-746-6
Perfect score: 2316
Sequence: 1 AVSTSPAPAFDTOTLIANAHA.....NPPLTALTAACCTPGARA 438

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgm2_6/ptodata/1/1aa/5_COMB.pep:*
2: /cgm2_6/ptodata/1/1aa/6_COMB.pep:*
3: /cgm2_6/ptodata/1/1aa/H_COMB.pep:*
4: /cgm2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
5: /cgm2_6/ptodata/1/1aa/RE_COMB.pep:*
6: /cgm2_6/ptodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2316	100.0	438	2	US-09-674-866A-6 Sequence 6, Appli
2	2316	100.0	1359	2	US-09-674-866A-3 Sequence 3, Appli
3	437.5	18.9	1236	2	US-10-023-649A-4 Sequence 4, Appli
4	437.5	18.9	1236	2	US-10-023-649A-6 Sequence 6, Appli
5	402.5	17.4	1253	1	US-07-920-281C-3 Sequence 3, Appli
6	402.5	17.4	1253	2	US-08-466-277C-3 Sequence 3, Appli
7	402.5	17.4	1253	1	US-08-688-842-3 Sequence 3, Appli
8	395.5	17.1	1245	1	US-08-801-263A-6 Sequence 6, Appli
9	395.5	17.1	1245	2	US-09-103-248-6 Sequence 6, Appli
10	395.5	17.1	1245	2	US-09-367-764-6 Sequence 6, Appli
11	394	17.0	981	2	US-09-991-181-355 Sequence 13, Appli
12	389.5	16.8	1245	1	US-08-801-263A-3 Sequence 3, Appli
13	389.5	16.8	1245	2	US-09-102-248-10 Sequence 10, Appli
14	389.5	16.8	1245	2	US-09-367-764-10 Sequence 10, Appli
15	383.5	16.6	1245	1	US-09-102-248-10 Sequence 10, Appli
16	383.5	16.6	1245	2	US-09-367-764-10 Sequence 10, Appli
17	383.5	16.6	1245	2	US-09-367-764-10 Sequence 10, Appli
18	113	4.9	2315	2	US-09-543-681A-5434 Sequence 5434, Ap
19	107	4.6	437	2	US-09-830-189C-2 Sequence 2, Appli
20	107	4.6	437	2	US-09-991-181-355 Sequence 355, App
21	107	4.6	437	2	US-09-990-444-355 Sequence 355, App
22	107	4.6	437	2	US-10-033-301-16 Sequence 16, Appli
23	107	4.6	437	2	US-09-997-333-355 Sequence 355, App
24	107	4.6	437	2	US-09-992-598-355 Sequence 355, App
25	106	4.6	437	2	US-09-073-569-2 Sequence 2, Appli
26	103.5	4.5	298	2	US-09-248-796A-26762 Sequence 26762, A
27	99.5	4.3	685	2	US-10-104-047-2316 Sequence 2316, Ap

28	99.5	4.3	985	2	US-09-993-777-6 Sequence 6, Appli
29	99.5	4.3	985	2	US-09-993-777-6 Sequence 6, Appli
30	99.5	4.3	985	4	PCT-US96-03916-6 Sequence 6, Appli
31	99.5	4.3	985	4	PCT-US96-03916-6 Sequence 6, Appli
32	99.5	4.3	503	2	US-09-487-558B-424 Sequence 424, App
33	98.5	4.3	495	2	US-09-949-016-7846 Sequence 7846, App
34	97.5	4.2	1125	2	US-09-900-920-60 Sequence 60, Appli
35	97.5	4.2	1398	2	US-09-543-681A-5641 Sequence 5641, Ap
36	96.5	4.2	1402	2	US-09-125-635-12 Sequence 12, Appli
37	96	4.1	1191	2	US-09-248-796A-16243 Sequence 16243, A
38	95	4.1	382	2	US-09-603-208A-40 Sequence 40, Appli
39	95	4.1	352	2	US-09-702-705-334 Sequence 334, App
40	95	4.1	582	2	US-09-736-457-334 Sequence 334, App
41	95	4.1	582	2	US-09-614-124B-334 Sequence 334, App
42	95	4.1	582	2	US-09-671-325-334 Sequence 334, App
43	95	4.1	582	2	US-09-589-184-334 Sequence 334, App
44	95	4.1	582	2	US-09-658-824-334 Sequence 334, App
45	95	4.1	582	2	US-10-017-754-334 Sequence 334, App

ALIGNMENTS

RESULT 1					
US-09-674-866A-6					
Sequence 6, Application US/09674866A					
Patent No. 6719980					
GENERAL INFORMATION:					
APPLICANT: Akzo No. 6719980el NV					
TITLE OF INVENTION: Structural Proteins of Fish Pancreatic Disease Virus					
FILE OF INVENTION: and Uses Thereof					
FILE REFERENCE: 1/98376US					
CURRENT APPLICATION NUMBER: US/09/674,866A					
CURRENT FILING DATE: 2002-06-17					
PRIOR APPLICATION NUMBER: EP98201461.5					
PRIOR FILING DATE: 1998-05-08					
NUMBER OF SEQ ID NOS: 15					
SOFTWARE: PatentIn Ver. 2.1					
SEQ ID NO 6					
LENGTH: 438					
TYPE: PRT					
ORGANISM: Salmon pancreatic disease virus					
FEATURE:					
OTHER INFORMATION: B2					
US-09-674-866A-6					
Query Match					
Best Local Similarity 100.0%; Score 2316; DB 2; Length 438;					
Matches 438; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	AVSTSPAPAFDTOTLIANAHAASPRAYCPDGDGACISPLAIDVSSGSDHVRMRVGS	60		
DB	1	AVSTSPAPAFDTOTLIANAHAASPRAYCPDGDGACISPLAIDVSSGSDHVRMRVGS	60		
QY	61	QSGVTANGGAAGETSLRYLGRDGKVAADNTRLVRTTAKCDVLAQNGHYLIANCPIVQS	120		
DB	61	QSGVTANGGAAGETSLRYLGRDGKVAADNTRLVRTTAKCDVLAQNGHYLIANCPIVQS	120		
QY	121	LTVAATLDGRHOCTTVFHHQVTEKTRFRSKGHLSDMTKCTRFSTTPKSAALYLVDV	180		
DB	121	LTVAATLDGRHOCTTVFHHQVTEKTRFRSKGHLSDMTKCTRFSTTPKSAALYLVDV	180		
QY	181	YDALPIVEISTVTCSDQCTVAVPPGTYVKKPKKKSADSAVTFTSDQCTCERPV	240		
DB	181	YDALPIVEISTVTCSDQCTVAVPPGTYVKKPKKKSADSAVTFTSDQCTCERPV	240		
QY	241	LTAASTIQGPHLSAALPGSGKEVKARIPFPPTATCRVSAVAPLSITYESDVLAA	300		
DB	241	LTAASTIQGPHLSAALPGSGKEVKARIPFPPTATCRVSAVAPLSITYESDVLAA	300		
QY	301	GTATYCVLLTTRNIGFHSNATSEMIQCKTLRIIVTPOGIELTGNNAEMHFMSSVRYAS	360		
DB	301	GTATYCVLLTTRNIGFHSNATSEMIQCKTLRIIVTPOGIELTGNNAEMHFMSSVRYAS	360		

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2006, 14:32:50 ; Search time 172 Seconds
(without alignments)
1064.007 Million cell updates/sec

Title: US-10-788-746-6

Perfect score: 2316

Sequence: 1 AVSTSPAAFYDTQILAAHAA.....NPPLTALTAACCPGARA 438

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:
1: /cgn2_6/ptodata/1/pubpa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpa/US08_PUBCOMB.pep.*
3: /cgn2_6/ptodata/1/pubpa/US09_PUBCOMB.pep.*
4: /cgn2_6/ptodata/1/pubpa/US10_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpa/US10B_PUBCOMB.pep.*
6: /cgn2_6/ptodata/1/pubpa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2316	100.0	438	5	US-10-788-746-6 Sequence 6, App1
2	2316	100.0	1359	5	US-10-788-746-3 Sequence 3, App1
3	446.5	19.3	487	5	US-10-811-353-8 Sequence 8, App1
4	441.5	18.1	1254	5	US-10-811-353-5 Sequence 5, App1
5	437.5	18.9	1236	4	US-10-023-649-4 Sequence 4, App1
6	437.5	18.9	1236	4	US-10-023-649-6 Sequence 6, App1
7	437.5	18.9	1236	5	US-10-921-868A-6 Sequence 6, App1
8	437.5	18.9	1236	5	US-10-921-868A-6 Sequence 6, App1
9	402.5	17.4	1253	3	US-09-901-106-3 Sequence 3, App1
10	394	17.0	981	3	US-09-991-258-13 Sequence 13, App1
11	394	17.0	981	5	US-10-929-234-13 Sequence 13, App1
12	121	5.2	594	4	US-10-437-963-173742 Sequence 29, App1
13	113	4.9	2000	4	US-10-010-901-29 Sequence 4, App1
14	113	4.9	4262	5	US-10-704-781-3 Sequence 4, App1
15	113	4.9	4493	5	US-10-704-781-3 Sequence 5927, App1
16	109.5	4.7	472	5	US-10-739-930-5927 Sequence 355, App1
17	107	4.6	437	3	US-09-989-722-355 Sequence 355, App1
18	107	4.6	437	3	US-09-989-722-355 Sequence 355, App1
19	107	4.6	437	3	US-09-989-722-355 Sequence 355, App1
20	107	4.6	437	3	US-09-989-722-355 Sequence 355, App1
21	107	4.6	437	3	US-09-989-722-355 Sequence 355, App1
22	107	4.6	437	3	US-09-989-722-355 Sequence 355, App1
23	107	4.6	437	3	US-09-991-073-355 Sequence 355, App1
24	107	4.6	437	3	US-09-990-442-355 Sequence 355, App1
25	107	4.6	437	3	US-09-991-163-355 Sequence 355, App1
26	107	4.6	437	3	US-09-993-604-355 Sequence 355, App1
27	107	4.6	437	3	US-09-990-456-355 Sequence 355, App1

28	107	4.6	437	3	US-09-989-721-355 Sequence 355, App1
29	107	4.6	437	3	US-09-992-598-355 Sequence 355, App1
30	107	4.6	437	3	US-09-989-293A-355 Sequence 355, App1
31	107	4.6	437	3	US-09-989-725-355 Sequence 355, App1
32	107	4.6	437	3	US-09-990-444-355 Sequence 355, App1
33	107	4.6	437	3	US-09-991-181-355 Sequence 355, App1
34	107	4.6	437	3	US-09-989-730-355 Sequence 355, App1
35	107	4.6	437	3	US-09-990-443-355 Sequence 355, App1
36	107	4.6	437	3	US-09-993-687-355 Sequence 355, App1
37	107	4.6	437	3	US-09-989-724-355 Sequence 355, App1
38	107	4.6	437	3	US-09-997-653-355 Sequence 355, App1
39	107	4.6	437	3	US-09-989-724-355 Sequence 355, App1
40	107	4.6	437	3	US-09-989-728-355 Sequence 355, App1
41	107	4.6	437	3	US-09-990-441-355 Sequence 355, App1
42	107	4.6	437	3	US-09-993-667-355 Sequence 355, App1
43	107	4.6	437	3	US-09-997-428-355 Sequence 355, App1
44	107	4.6	437	3	US-09-997-666-355 Sequence 355, App1
45	107	4.6	437	3	US-09-990-438-355 Sequence 355, App1

ALIGNMENTS

RESULT 1					
US-10-788-746-6					
Sequence 6, Application US/10788746					
Publication No. US20040258707A1					
GENERAL INFORMATION:					
APPLICANT: Akzo Nobel NV					
TITLE OF INVENTION: Structural Proteins of Fish Pancreatic Disease Virus					
FILE REFERENCE: I/98376US					
CURRENT APPLICATION NUMBER: US/10/788,746					
CURRENT FILING DATE: 2004-02-26					
PRIOR APPLICATION NUMBER: US/09/674,866					
PRIOR FILING DATE: 2001-06-04					
PRIOR APPLICATION NUMBER: EP98201461.5					
PRIOR FILING DATE: 1998-05-08					
NUMBER OF SEQ ID NOS: 15					
SOFTWARE: PatentIn Ver. 2.1					
SEQ ID NO 6					
LENGTH: 438					
TYPE: PRT					
ORGANISM: Salmon pancreatic disease virus					
FEATURE:					
OTHER INFORMATION: E2					
US-10-788-746-6					
Query Match					
Best Local Similarity 100.0%; Score 2316; DB 5; Length 438;					
Matches 438; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1	AVSTSPAAFYDTQILAAHAAAPYRACPCDGTACISPIAIDEVSSGSDHLMRVGS	60		
DB	1	AVSTSPAAFYDTQILAAHAAAPYRACPCDGTACISPIAIDEVSSGSDHLMRVGS	60		
QY	61	OSGTAKGGAAGFSLRLGDDGKHADNRLVYRTAKCDVLAOTGHYILANCPVGS	120		
DB	61	OSGTAKGGAAGFSLRLGDDGKHADNRLVYRTAKCDVLAOTGHYILANCPVGS	120		
QY	121	LTVATLPGTHOCTVEHGVTEKTRERKSHLSMTKCTKTFSTPKKALYLDV	180		
DB	121	LTVATLPGTHOCTVEHGVTEKTRERKSHLSMTKCTKTFSTPKKALYLDV	180		
QY	181	YDALPISVETVVTGSDQCTVRVPETVYKPKCKSADSAITFTSDQTFCEBPV	240		
DB	181	YDALPISVETVVTGSDQCTVRVPETVYKPKCKSADSAITFTSDQTFCEBPV	240		
QY	241	LTAASITGKPHLSAMPSGKGVKARIIPPEPETAICVSVAPLSITTEESDVLAA	300		
DB	241	LTAASITGKPHLSAMPSGKGVKARIIPPEPETAICVSVAPLSITTEESDVLAA	300		
QY	301	GTAKPVALLTRNLGFHSNATSEWIOGKYLRIIPVTPGIBLTWGNAPMHWSSVRYAS	360		

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2006, 14:33:40 ; Search time 27 Seconds

(without alignments)
713.825 Million cell updates/sec

Title: US-10-788-746-6

Perfect score: 2316
Sequence: 1 AVSTSPAFYDTQILAAHAA.....NPPLTALTAACCTPGARA 438

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 225428 seqs, 44002918 residues

Total number of hits satisfying chosen parameters: 225428

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database:

Published Applications AA New:
1: /SIDS5/prodata/1/pubppa/US08_NEW_PUB.pep.*
2: /SIDS5/prodata/1/pubppa/US06_NEW_PUB.pep.*
3: /SIDS5/prodata/1/pubppa/US07_NEW_PUB.pep.*
4: /SIDS5/prodata/1/pubppa/US09_NEW_PUB.pep.*
5: /SIDS5/prodata/1/pubppa/US10_NEW_PUB.pep.*
6: /SIDS5/prodata/1/pubppa/US11_NEW_PUB.pep.*
7: /SIDS5/prodata/1/pubppa/US12_NEW_PUB.pep.*
8: /SIDS5/prodata/1/pubppa/US60_NEW_PUB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	107.5	4.6	5291	US-11-052-554A-281	Sequence 281, App
2	107	4.6	437	US-10-131-826A-466	Sequence 466, App
3	107	4.6	437	US-10-973-115B-466	Sequence 466, App
4	107	4.6	437	US-10-213-535-16	Sequence 16, App
5	107	4.6	437	US-10-218-784-162	Sequence 162, App
6	107	4.6	437	US-10-219-061-162	Sequence 162, App
7	107	4.6	437	US-10-219-062-162	Sequence 162, App
8	107	4.6	437	US-10-219-064-162	Sequence 162, App
9	107	4.6	437	US-10-233-134-162	Sequence 162, App
10	107	4.6	437	US-11-290-153-466	Sequence 466, App
11	99.5	4.3	685	US-11-072-512-2916	Sequence 2916, App
12	98.5	4.3	487	US-10-745-586-11	Sequence 11, App
13	97	4.2	403	US-10-453-372-1170	Sequence 1170, App
14	96	4.1	381	US-11-087-099-3465	Sequence 3465, App
15	96	4.1	940	US-11-045-004-39	Sequence 39, App
16	95	4.1	382	US-10-703-799B-40	Sequence 40, App
17	95	4.1	1709	US-10-995-561-973	Sequence 973, App
18	95	4.1	1709	US-10-453-372-410	Sequence 410, App
19	94.5	4.1	334	US-11-096-568A-11273	Sequence 11273, App
20	94.5	4.1	365	US-11-096-568A-11272	Sequence 11272, App
21	94.5	4.1	419	US-11-096-568A-11271	Sequence 11271, App
22	94.5	4.1	583	US-11-080-991-64	Sequence 64, App
23	94	4.1	626	US-11-045-004-25	Sequence 25, App
24	94	4.1	1700	US-10-453-372-398	Sequence 398, App
25	94	4.1	1700	US-10-453-372-412	Sequence 412, App

26	94	4.1	1700	US-10-453-372-414	Sequence 414, App
27	94	4.1	1700	US-10-453-372-418	Sequence 418, App
28	93.5	4.0	859	US-11-188-743-16	Sequence 16, App
29	93.5	4.0	859	US-11-007-428-5	Sequence 5, App
30	93.5	4.0	859	US-11-183-294-14	Sequence 14, App
31	93.5	4.0	859	US-11-235-037-7	Sequence 7, App
32	93	4.0	701	US-11-188-298-13650	Sequence 13650, App
33	93	4.0	1700	US-10-453-372-416	Sequence 416, App
34	92.5	4.0	3132	US-11-087-099-1245	Sequence 1245, App
35	92	4.0	1075	US-11-100-640-12	Sequence 12, App
36	92	4.0	1709	US-11-135-855-35	Sequence 35, App
37	91.5	4.0	347	US-11-221-203-12	Sequence 12, App
38	91.5	4.0	799	US-10-493-909-8	Sequence 8, App
39	91.5	4.0	822	US-10-493-909-48	Sequence 48, App
40	91.5	4.0	1571	US-11-052-554A-2	Sequence 2, App
41	91	3.9	1694	US-11-135-855-136	Sequence 36, App
42	90	3.9	934	US-10-453-372-1158	Sequence 1158, App
43	89.5	3.9	1461	US-11-052-554A-283	Sequence 283, App
44	89	3.8	619	US-10-999-886-3	Sequence 3, App
45	89	3.8	619	US-11-136-244-3	Sequence 3, App

ALIGNMENTS

RESULT 1	US-11-052-554A-281	Sequence 281, Application US/11052554A	Sequence 414, App
Publication No.	US2005028866A1		Sequence 418, App
GENERAL INFORMATION:			Sequence 16, App
APPLICANT:	Sachdeva, et al.		Sequence 5, App
TITLE OF INVENTION:	COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE		Sequence 14, App
FILE REFERENCE:	30853/40359A		Sequence 7, App
CURRENT APPLICATION NUMBER:	US/11/052,554A		Sequence 13650, App
CURRENT FILING DATE:	2005-02-07		Sequence 416, App
PRIOR APPLICATION NUMBER:	US 60/589,227		Sequence 1245, App
PRIOR FILING DATE:	2004-07-20		Sequence 12, App
PRIOR APPLICATION NUMBER:	IN 173/DEL/2004		Sequence 35, App
PRIOR FILING DATE:	2004-02-06		Sequence 12, App
NUMBER OF SEQ ID NOS:	763		Sequence 8, App
SOFTWARE:	Patent in version 3.3		Sequence 48, App
SEQ ID NO 281			Sequence 2, App
LENGTH:	5291		Sequence 36, App
TYPE:	PRT		Sequence 1158, App
ORGANISM:	Escherichia coli 0157:H7		Sequence 283, App
US-11-052-554A-281			Sequence 3, App
Query Match	4.6%; Score 107.5; DB 7; Length 5291;		
Best Local Similarity	19.8%; Pred. No. 1.7;		
Matches 104; Conservative	68; Mismatches 175; Indels 177; Gaps 22;		
2 VSTSPAFYDTQILAAHAAAPRAYCP---DCDGTACISPIAIDVSSGSDHVLNR 57			
1902 VSANPAGSVTTAAAGSTAGNPSVTPVTLNAAVAINATADVDVINAKEGALTL 1961			
58 VGSQSGVTAK-----CGAGETSLRYLG-----RDGKHA----- 87			
1962 SGSTSGVAGAGTATTTCGGKTSATVANGSWSTVPAADMAALRDGASASQASVNVNG 2021			
88 -----ADNRLVVRTAKCDVLAQ-----TGHYILANCPVGSULTVAAT 126			
2022 NSATTTHAVSDASAPYTTINTIAGDDITLNAABGALITIGS-----STAEAGQTVTV--T 2076			
127 LDGTRHQQTT-----VFEHOYTERKTERSKGHTLSMTKTC 163			
2077 LNTGYTGTVDGDSWSVSPSADLSTLTANNTVNAVSPKAGNPASVNHLLT----- 2130			
164 TRFSTTPKSKALVVDVADALPISVEISTV-----VTCSDSQCTVAVPPKTVKPKKCS 219			
2131 -----VDTSVPV-VTINTVAGDDVYNAHQAQITIGSAT-----GA 2167			
220 ADSATVTFSDSQFT-----CREPYLTASITQGRHLASAMLPSSGKEVKARI 269			

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2006, 14:28:04 ; Search time 42 Seconds
(without alignments)
1003.403 Million cell updates/sec

Title: US-10-788-746-6

Perfect score: 2316
Sequence: 1 AVGSTPAFDTQTLAAHAA.....NPPPLTALTAACCTPGARA 438

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 80: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	456.5	19.7	1241	2 S26373	genome polypeptide
2	454.5	19.6	1240	1 VHWVWV	structural polypeptide
3	452.5	19.5	1242	2 S72350	structural polypeptide
4	452	19.5	1247	1 VHWVWV	structural polypeptide
5	446.5	19.3	422	1 VHWVWV	structural polypeptide
6	442.5	19.1	1242	2 A56605	structural polypeptide
7	437.5	18.9	1239	1 VHWVWV	structural polypeptide
8	434.5	18.8	1236	1 VHWVWV	structural polypeptide
9	416.5	18.0	1254	1 VHWVWV	structural polypeptide
10	413.5	17.9	422	1 VHWVWV	structural polypeptide
11	399	17.2	1254	1 JQ1978	structural polypeptide
12	398.5	17.2	1253	1 VHWVWV	structural polypeptide
13	397.5	17.2	1245	1 VHWVWV	structural polypeptide
14	396	17.1	1254	1 JQ1979	structural polypeptide
15	396	17.1	1255	1 D44213	structural polypeptide
16	395	17.1	1254	1 VHWVWV	structural polypeptide
17	394.5	17.0	755	2 S42462	structural polypeptide
18	393.5	17.0	1145	2 S37136	structural polypeptide
19	393	17.0	1254	1 VHWVWV	structural polypeptide
20	386.5	16.7	1245	1 VHWVWV	structural polypeptide
21	386.5	16.7	1245	1 VHWVWV	structural polypeptide
22	384	16.6	1255	1 B44213	structural polypeptide
23	383.5	16.6	423	1 VHWVWV	structural polypeptide
24	126	5.4	69	2 G37264	E2 glycoprotein -
25	115	5.0	69	2 G37264	E2 glycoprotein -
26	109.5	4.7	472	2 T05852	hypothetical prote
27	109	4.7	1777	2 T34369	hypothetical prote
28	107.5	4.6	5188	2 B85547	probable RTX faml
29	107.5	4.6	5291	2 F90696	hypothetical prote

30	106.5	4.6	1251	2 T21389	hypothetical prote
31	104.5	4.5	1609	2 S25345	probable membrane
32	103.5	4.5	551	2 G84301	hypothetical prote
33	103	4.4	1694	2 S50065	staiadhesin - mou
34	102.5	4.4	967	2 S66852	hypothetical prote
35	101.5	4.4	1522	2 H88380	protein T2287.3 [i
36	101	4.4	587	2 JH0464	DM-GRASP precursor
37	101	4.4	588	2 JH0506	adhesion molecule
38	99	4.3	503	2 S63257	probable membrane
39	99	4.3	630	2 A39344	tumor-associated m
40	98.5	4.3	487	2 JCT126	testis zinc finger
41	98	4.2	588	2 A45254	surface glycoprote
42	97	4.2	66	2 A37264	E2 glycoprotein -
43	97	4.2	1676	2 A56508	anucleate primary
44	96.5	4.2	747	2 S18685	Sc/SwM protein -
45	96.5	4.2	748	2 S18690	Sc/SvQ protein - E

ALIGNMENTS

RESULT 1

S26373 genome polypeptide - eastern equine encephalomyelitis virus
N:Contains: 6X protein; capsid protein C; envelope protein E1; envelope protein E2; en
C:Species: eastern equine encephalomyelitis virus
C>Date: 06-Jan-1994 #sequence_revision 01-Nov-1996 #text_change 31-Dec-2004
C/Accession: S26373
R:Volchikov, V.B.; Volchikova, V.A.; Netesov, S.V.
Mol. Gen. Microbiol. Virusol. 5, 8-15, 1991
A>Title: Complete nucleotide sequence of the eastern equine encephalomyelitis virus ge
A:Reference number: S26369; MUID:91375524; PMID:1896061
A/Accession: S26373
A/Molecule type: mRNA
A/Residues: 1-1241 <VOL>
A/Cross-references: UNIPROT:Q66579; UNIPARC:UP100000F1639; EMBL:X63135; NID:959185; PI
A/Note: sequence could not be checked because of bad print in paper
C/Superfamily: pestivirus genome polypeptide
C/Keywords: capsid protein, envelope protein, glycoprotein, polypeptide
F/1-260/Product: capsid protein C #status predicted <CNP>
F/261-323/Product: envelope protein E1 #status predicted <EP1>
F/324-743/Product: envelope protein E2 #status predicted <EP2>
F/744-800/Product: 6X protein #status predicted <6XP>
F/801-1241/Product: envelope protein E1 #status predicted <EP1>

Query Match	Score	DB 2	Length	ID
Query Match	19.7%	Score 456.5	DB 2	Length 1241
Best Local Similarity	30.2%	Pred. No. 5.3e-27		
Matches	133	Conservative	64	Mismatches 189; Indels 55; Gaps 17
QY	21	ASPRAYCPDGDGACTSPALIDRVSSGSDHVRMRVSGSGVYAKGAGTSLRYL- 79		
DB	335	ARPIADPCPNCGHRCSPALIBSRGDANAGVIRIQTSNFGAKTDD- -VDLAYMS 389		
QY	80	---GRDGVNAADTRTLVVTAKCDVLOATGHTLANCPVGSLLVVAATLDGRHOCT 136		
DB	390	FMNGKTKSKIKIDN--LHVRTSAPCSLVHHGYIILAQCPGDIIVTVGFHDGPRNHTCV 447		
QY	137	VFEHQV-----TEKTRRSKGNHLSDMTKCTRFSTTPPKSALYLVVDYDALPISVEI 190		
DB	448	A--HKVEFRPVGRKRYRHPREHGVEL-----PCNRY-THKADQGHVVEHQ--FGLVAD 497		
QY	191	STVTCGSDQCVAVVPPTGTYKPDCKKSADSAVTPFSDSQGTCEFPVLTAAISITGK 250		
DB	498	HSLLSHSAKVKITVPSGAQVKYCKCPDVBEGT--TSSDYTTICDVQCKRAVLINLK 554		
QY	251	PHL-PSAMLPSSG-SKEVVARIPFPPEPETAICRVSAVLPISITVRESVLLAGTAKYVPL 308		
DB	555	KVAVNSGRILPRGSDTRKGLHVFVYKAKCIATLADPELVKHKRTLILHLVPDHTL 614		
QY	309	LTNNGHSHNSATSEMTIGKTLRIIPVTPGIEILTWGNANPMHFWSSRYVAGDAAYPW 368		
DB	615	LTTRSLGSDANPTQWIERPTTVNFTVGTGBLBYTWGNHHPKRWAAQ--ESGEGNPHGWPH 673		